

Modelling Infrastructure Systems for Resilience and Sustainability

Sarah Dunn^{a*}

Sean Wilkinson^a

Gaihua Fu^a

Richard Dawson^a

Abstract: Modern infrastructure systems are vital to the functioning of modern society. They promote social well-being, support economic development and are crucial in mitigating the effects of natural hazards. While there is some understanding of their mitigation role, there has been little quantifiable work on how they support our societies or how they stimulate economic development. Some recent analysis of infrastructure systems have shown that many of these seemingly different systems display similar architectures to each other leading to the hypothesis that the evolution of these systems is a result of underlying drivers that are common to all. This paper presents a network model that captures the growth of infrastructure networks in terms of architecture, hazard tolerance and geographical characteristics. The results presented in the paper suggest that the model may be the basis for an enhanced understanding of the role that infrastructure plays in sustaining our communities.

Key words: Complex networks; Graph theory; Infrastructure systems; Modelling; Resilience.

I. Introduction

Infrastructure systems, such as water, transport, communication and energy networks form the backbone of our modern communities¹ and are crucial to the functioning of our modern society². These systems not only promote social well-being and support economic productivity, but they also play a crucial role in mitigating the effects of natural hazards and as such are designed to have lower than usual probabilities of failure. However, it is not possible to engineer out all possibility of failure and therefore not only does failure occur, but when it does it can lead to great suffering in our communities. For example, in the aftermath of hurricane Katrina two dozen hospitals were left without electricity, meaning that they could not operate essential laboratory and x-ray equipment, dialysis machines and ventilators, resulting in many potentially preventable deaths³. The above argument demonstrates not only the need to ensure that these systems can provide at least a baseline level of service in the aftermath of a disaster, but also to develop tools that can assess the behaviour of these systems subjected to very rare events and what this baseline should be.

The current design approach for our infrastructure systems is to design the individual components to have a particular probability of failure. However, this silo-based approach

^a School of Civil Engineering and Geosciences, Newcastle University, Newcastle, NE1 7RU, UK.

*Corresponding author: Sarah Dunn sarah.dunn@ncl.ac.uk
<http://dx.doi.org/10.14453/isngi2013.proc.17>

does not ensure adequate system performance in the overload situation (i.e. it does not guarantee that a baseline level of service will be delivered when these systems become damaged by the effects of a natural disaster). Systems are traditionally modelled using physically based models (e.g. a hydraulic model for a water distribution system), which are useful at providing scenario based information. However, due to their complexity, they can be found lacking when used to inform us of the resilience of the system and highlight structural inadequacies.

To solve this problem, recent studies have applied network graph theory to model the complex interactions between individual components. In this approach, only the topology of the system is considered which is modelled as a series of nodes and connecting links. Studies using this approach have discovered that infrastructure systems display surprising similar network architectures, which has led to the hypothesis that the evolution of these systems is a result of underlying drivers that are common to all. However, this approach can be deficient when used to model systems spread over wide geographic regions as it purely consider the connections between components and not the locations of the components themselves. In this paper, we present a network model which captures these underlying drivers governing connections between individual components, but extend this to capture the drivers governing the geographic growth of these systems.

II. Modelling Infrastructure Systems using Topological Complex Networks

Infrastructure systems can be modelled using network graph theory by using nodes to represent the individual components (e.g. power stations, communities in an electrical distribution system for example) and links to model the connections between these individual components (e.g. the transmission lines). Previous studies analysing infrastructure systems have shown that many of these systems naturally configure to specific network architectures or network classes. These network classes describe different patterns of nodal connectivity (i.e. different arrangements of the nodes and links) and are characterised by a 'degree distribution'. It is this distribution that allows for the distinction between different classes of network. The degree distribution of a network is the probability distribution of the degrees of all nodes in the network, where the degree of a node is equal to the number of links connected to it. Infrastructure systems have been shown to belong to one of two network classes: the scale-free or exponential network class^{4,5,6}. Both of these classes are similar in that they comprise a small number of highly connected nodes and a larger number of poorly connected nodes. A sample scale-free network and a characteristic degree distribution are shown in Figure 1.

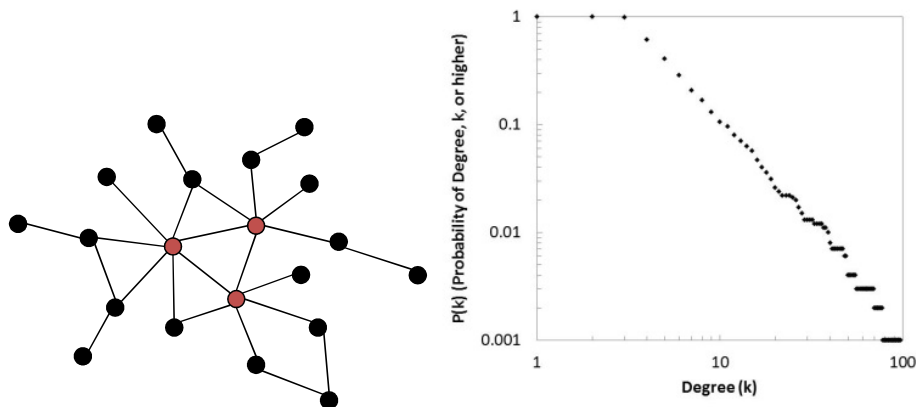


Figure 1. Showing (a) a sample scale-free network and (b) the characteristic degree distribution of a scale-free network.

The Internet and the World-Wide-Web were the first infrastructure systems to be classified as scale-free by Albert et al.⁴. In a further study they went on to consider the underlying drivers which governed the formation of connections between nodes in these networks and used this to develop a network generation algorithm capable of forming synthetic networks with the scale-free degree distribution⁷. This algorithm is based upon the ideas of *growth* and *preferential attachment*⁸ and starts with an initial number of isolated nodes, m_0 , which are usually a small percentage of the total number of nodes in the network. New nodes are then added to the network at each ‘time-step’ (incorporating the idea of *growth*) until the total number of nodes in the network is reached. These added nodes have between 1 and m_0 links attached to them and connect to the existing nodes in the network based upon the idea of *preferential attachment*. As such, the probability of connecting to an existing node is based upon the degree, with nodes with a higher degree being more likely to ‘attract’ a link from the new node (i.e. the rich get richer). It is this rule of *preferential attachment* which results in the formation of a few highly connected nodes and many poorly connected nodes in the network (as seen in Figure 1).

Other infrastructure systems have been classified as exponential networks, the most notable of which are power grids^{5,9,10,11}. In similar manner to scale-free networks, previous studies have also considered the drivers governing the formation of connections in this network class and have used this information to develop network generation algorithms. Wilkinson et al.¹² developed one such generation algorithm which was based upon the scale-free algorithm⁷, but with one notable difference. In their algorithm, Wilkinson et al.¹² also considered the geographical characteristics of the system, rather than focusing on topology alone. Many infrastructure systems which belong to the exponential network class are distributed over wide geographic areas, for example power grids can be distributed over whole countries or even continents. To reflect this in their algorithm, Wilkinson et al.¹² modified the idea of *preferential attachment* to incorporate a spatial component. They proposed that poorly connected nodes can capitalise on their close proximity to a highly connected node by attracting links that were bound for the high degree node, arguing that the probability of attachment is based upon both degree and proximity. They showed that this network generation algorithm is capable of generating proxy networks for real world infrastructure systems (using the European air traffic network as an example).

One of the main advantages of classifying infrastructure systems (other than identifying the underlying drivers) is that this information can be used to gain an insight into the inherent hazard tolerance of each system. For example, both scale-free and exponential networks have been shown to be resilient to a random hazard (e.g. failure due to random events, such as lack of maintenance), but vulnerable to targeted attack (e.g. a targeted terrorist attack). This is because a random hazard has a small chance of removing a highly connected node, whereas a targeted attack will often remove these important nodes seeking to cause the maximum disruption to the network¹³. It can be seen from Figure 1(a) that the removal of one of the three high degree nodes (red) will cause a larger impact to the network than the removal of one of the more numerous smaller degree nodes (black).

III. Development of Spatial Network Model

Whilst, there has been a great deal of research considering the topological hazard tolerance of these networks, there has currently been very little research on the implications of the geographical distribution of an infrastructure system on its hazard tolerance. Traditional

graph theory focuses on topological models where it is only the presence of a connection between two nodes which is considered to be important. The little work that has studied real world spatial networks focuses mainly on characterising the topology of the system (into one of the network classes), while the spatial element of the same network receives less attention - if not neglected entirely⁸. This spatial component may not seem important; however as shown in the development of the exponential generation algorithm and subsequent hazard tolerance analysis by Wilkinson et al.¹² it can not only have a significant influence on the layout of the network, but is crucial in defining its tolerance to real world hazards.

Traditional graph theory has discovered underlying drivers governing the formation of connections between pairs of nodes; we now expand upon this to consider the drivers determining the geographic location of nodes within a network. To achieve this we develop an algorithm to generate nodal layouts (with different characteristics) by simplifying the method of cellular automata. This technique has previously been used to predict urban growth around cities, including San Francisco¹⁴ and Washington¹⁵. However, these models require the input of detailed historical data (including: historical maps, aerial photographs and digital maps) at regular time intervals from the initial settlement in the study area to the present day. This data is normally obtained from a variety of sources and there are often problems with assembling the dataset, including: inconsistent dimensions of features, generalisation in historical maps, different projections of the study area and different coordinate systems. As such the main disadvantage in this method is that the accuracy of the results is highly dependent on the quality and quantity of the historical data¹⁵, indeed if insufficient historical data for the study area is obtained then it is not possible to generate a synthetic model. Therefore, we propose a new method which incorporates and simplifies the ideas behind cellular automata to generate proxy nodal layouts for real world networks.

In a similar manner to cellular automata, the algorithm requires the input of a set of initial conditions, from which the nodal layout forms over a given timeframe. These initial conditions define the spatial boundary of the network (or the study area), the number of seed nodes and the location and initial radius of these seed nodes. Using these inputs the network is allowed to 'grow' and the remaining nodes are added individually to the network at each 'time-step' until the total number of nodes is reached.

At each time-step the algorithm determines if an added node will be located within the radius of one of the individual clusters or will be located outside the influence of all of the clusters, depending on a user specified probability value. By allowing a small proportion of the total number of nodes in the network to be located outside the cluster radii, a rural environment over the whole of the spatial boundary is represented. However, if the added node is to be located inside the radius of a cluster, then this node is 'attracted' to the different individual clusters based upon a calculated probability value. This probability value is dependent upon the density of the cluster and is calculated by dividing the number of nodes within the influence of the cluster by the cluster radius. This probability value encompasses the idea that a city, with a high population density, can be expected to have more nodes (representing train stations, for example) than a rural community which has a significantly lower population density. With the addition of a new node to the cluster, the radius of the cluster is allowed to expand outwards, in order to simulate the 'growth' of an urban area. This growth is logarithmic meaning that it increases rapidly for the first few nodes added to the cluster and then reduces to only increasing marginally with further added nodes.

The algorithm also incorporates a variable to alter the density of the network as a whole; the visual effects of this variable (termed C_D) can be seen in Figure 2. These three nodal layouts have been generated using the same seed locations and radius values, but different C_D values (100, 200 and 300 respectively).

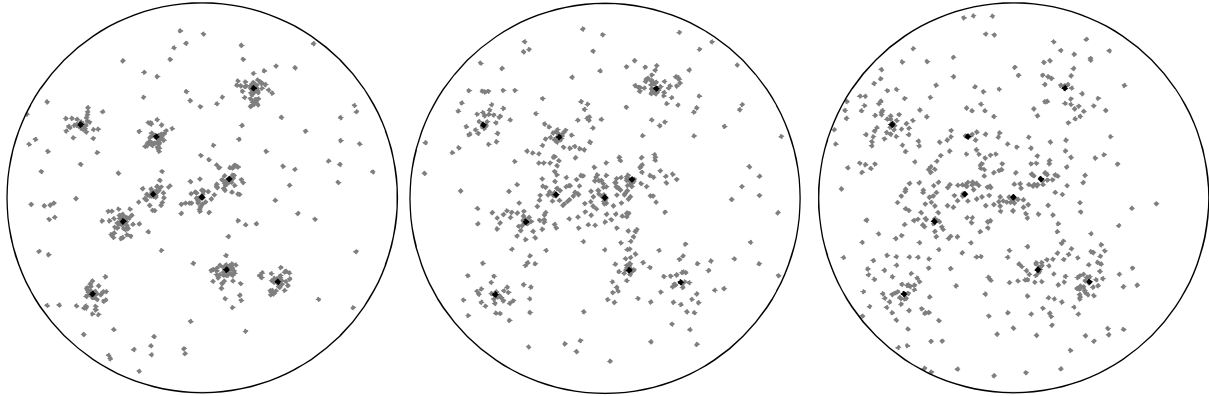


Figure 2. Three nodal layouts generated using the clustering algorithm with C_D values of (a) 100, (b) 200 and (c) 300. Where the black dots represent the seed nodes and the grey dots show the added nodes, the outer circle defines the spatial boundary of the network.

To validate this clustering algorithm we use it to generate a real world nodal layout for the Target dataset (shown in Figure 3 and obtained from edigitalz.COM¹⁶). In a similar manner to Wilkinson et al.¹² we also determine the spatial distribution of stores (nodes) in this network (shown as the black dots in Figure 3(b)). The proxy nodal layout for this network was generated by visually defining the location of the seed nodes (as the centre of the high density areas on the GIS image, Figure 3(a)) and determining their radius values by considering the density of the cluster. This information was then inputted into the clustering algorithm and nodal layouts generated for different C_D values (to find the best fit for the real world data). The spatial distribution for the proxy network has been shown as grey dots on the corresponding spatial distribution in Figure 3(b). It can be seen that this proxy nodal layout is in good agreement with the actual real world nodal layout, meaning that the model is capable of determining the underlying drivers governing the formation of real world nodal layouts. We can now use this information to assess the hazard tolerance of these spatial networks, as a geographic hazard will affect the nodes and links depending on their location.

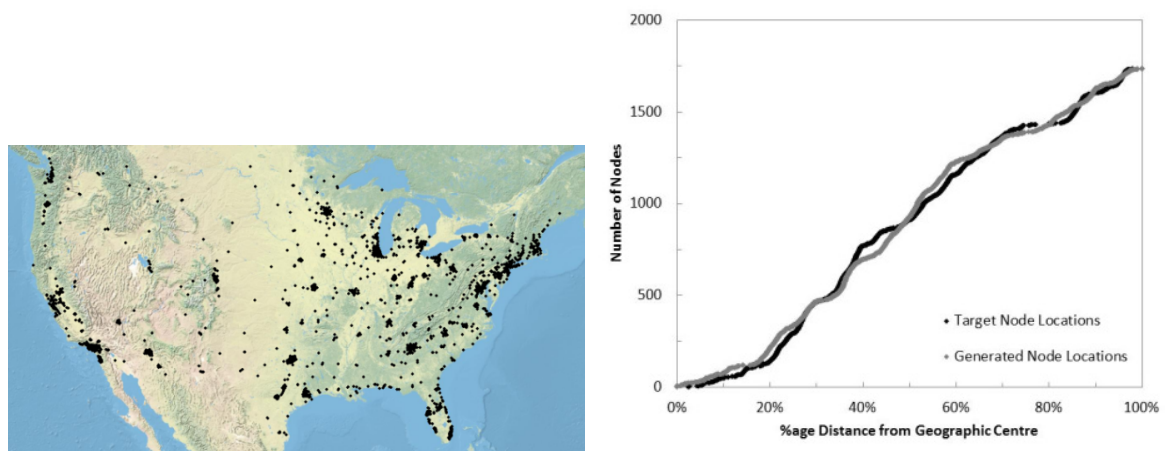


Figure 3. Showing (a) the spatial layout of nodes in the Target dataset and (b) the spatial distributions, where the black dots correspond to the actual network and the grey dots to the generated nodal layout.

IV. Conclusion

This paper has presented a network model which captures the growth of infrastructure systems in terms of their network architecture, geographical characteristics and hazard tolerance. Previous studies have considered the underlying drivers which determine the architecture of the network, but have failed to consider the spatial distribution of the same network, which can have dramatic effects to the perceived hazard tolerance of the network. Therefore, this paper developed a network model which also considers the underlying drivers governing the location of individual components within an infrastructure system. The results presented in the paper suggest that these models may form the basis for an enhanced understanding of the role that infrastructure plays in sustaining our communities and also in stimulating economic development.

References

- ¹Institution of Civil Engineers, *The State of the Nation: Defending Critical Infrastructure*, 2009.
- ²Murray, A., and Grubestic, T., *Critical Infrastructure: Reliability and Vulnerability*, Springer, 2007.
- ³Gray, B. H., and Herbert, K., *After Katrina: Hospitals in Hurricane Katrina - Challenges Facing Custodial Institutions in a Disaster*, 2006.
- ⁴Albert, R., Jeong, H., and Barabasi, A. L., "Internet - Diameter of the World-Wide Web", *Nature*, Vol. 401, No. 6749, 1999, pp. 130-131. <http://dx.doi.org/10.1038/43601>
- ⁵Crucitti, P., Latora, V., and Marchiori, M., "A topological analysis of the Italian electric power grid", *Physica a-Statistical Mechanics and Its Applications*, Vol. 338, No. 1-2, 2004, pp. 92-97.
- ⁶Costa, L. F., Oliveira, O. N., Travieso, G., Rodrigues, F. A., Boas, P. V., Antiqueira, L., Viana, M., and Rocha, L. E. C. d., "Analyzing and Modeling Real-World Phenomena with Complex Networks: A Survey of Applications", *Advances in Physics*, Vol. 60, No. 3, 2011, pp. 329-412. <http://dx.doi.org/10.1080/00018732.2011.572452>
- ⁷Barabasi, A. L., and Albert, R., "Emergence of scaling in random networks", *Science*, Vol. 286, No. 5439, 1999, pp. 509-512. <http://dx.doi.org/10.1126/science.286.5439.509>
- ⁸Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U., "Complex networks: Structure and dynamics", *Physics Reports-Review Section of Physics Letters*, Vol. 424, No. 4-5, 2006, pp. 175-308.
- ⁹Kinney, R., Crucitti, P., Albert, R., and Latora, V., "Modeling cascading failures in the North American power grid", *The European Physical Journal B - Condensed Matter and Complex Systems*, Vol. 46, No. 1, 2005, pp. 101-107.
- ¹⁰Rosas-Casals, M., Valverde, S., and Sole, R. V., "Topological Vulnerability of the European Power Grid under Errors and Attacks", *International Journal of Bifurcation and Chaos*, Vol. 17, No. 7, 2006, pp. 2465-2475. <http://dx.doi.org/10.1142/S0218127407018531>
- ¹¹Sole, R. V., Rosas-Casals, M., Corominas-Murtra, B., and Valverde, S., "Robustness of the European power grids under intentional attack." *Physical Review E*, Vol. 77, No. 2, 2008. <http://dx.doi.org/10.1103/PhysRevE.77.026102>
- ¹²Wilkinson, S., Dunn, S., and Ma, S., "The vulnerability of the European air traffic network to spatial hazards", *Natural Hazards*, Vol. 60, No. 3, 2012, pp. 1027-1036. <http://dx.doi.org/10.1007/s11069-011-9885-6>

- ¹³Albert, R., Jeong, H., and Barabasi, A. L., "Error and Attack Tolerance of Complex Networks", *Nature*, Vol. 406, No. 6794, 2000, pp. 378-382.
<http://dx.doi.org/10.1038/35019019>
- ¹⁴Clarke, K. C., Hoppen, S., and Gaydos, L., "A self-modifying cellular automation model of historical urbanisation in the San Francisco Bay area", *Environment and Planning B-Planning & Design*, Vol. 24, 1995, pp. 27-261.
- ¹⁵Clarke, K. C., and Gaydos, L., "Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore", *International Journal of Geographical Information Science*, Vol. 12, No. 7, 1998, pp. 699-714. <http://dx.doi.org/10.1080/136588198241617>
- ¹⁶edigitalz.COM, *World Digital Data*, 2012, URL: <http://www.edigitalz.com/>.